Approximation norms and duality for communication complexity lower bounds

Troy Lee
Columbia University

Adi Shraibman
Weizmann Institute
From min to max

- The cost of a “best” algorithm is naturally phrased as a minimization problem
- Dealing with this universal quantifier is one of the main challenges for lower bounders
- Norm based framework for showing communication complexity lower bounds
- Duality allows one to obtain lower bound expressions formulated as maximization problems
Communication complexity

• Two parties Alice and Bob wish to evaluate a function $f : X \times Y \to \{-1, +1\}$ where Alice holds $x \in X$ and Bob $y \in Y$.

• How much communication is needed?
Communication complexity

• Two parties Alice and Bob wish to evaluate a function \( f : X \times Y \rightarrow \{-1, +1\} \) where Alice holds \( x \in X \) and Bob \( y \in Y \).

• How much communication is needed? Can consider both deterministic \( D(f) \) and randomized \( R_\epsilon(f) \) versions.
Communication complexity

• Two parties Alice and Bob wish to evaluate a function $f : X \times Y \to \{-1, +1\}$ where Alice holds $x \in X$ and Bob $y \in Y$.

• How much communication is needed? Can consider both deterministic $D(f)$ and randomized $R_\epsilon(f)$ versions.

• Often convenient to work with communication matrix $A_f[x, y] = f(x, y)$. Allows tools from linear algebra to be applied.
How a protocol partitions communication matrix

Alice
X
Bob
Y
0
1

Bob
Y

Alice
X

0

1
How a protocol partitions communication matrix

<table>
<thead>
<tr>
<th>Alice X</th>
<th>Bob Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0</td>
<td>0 1</td>
</tr>
<tr>
<td>1 1</td>
<td>1 0</td>
</tr>
</tbody>
</table>
How a protocol partitions communication matrix

<table>
<thead>
<tr>
<th>Alice X</th>
<th>Bob Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>010</td>
</tr>
<tr>
<td>000</td>
<td></td>
</tr>
<tr>
<td>111</td>
<td>110</td>
</tr>
<tr>
<td>111</td>
<td></td>
</tr>
</tbody>
</table>

(Assuming X represents 0 and Y represents 1 for simplicity)
From min to max: Yao’s principle

- One of the best known examples of the min to max idea is Yao’s minimax principle:

\[ R_\epsilon(f) = \max_{\mu} D_\mu(f) \]

- To show lower bounds on randomized communication complexity, suffices to exhibit a hard distribution for deterministic protocols.

- The first step in many randomized lower bounds.
Let $A$ be a matrix. The singular values of $A$ are $\sigma_i(A) = \sqrt{\lambda_i(A^T A)}$.

Define

$$\|A\|_p = \ell_p(\sigma) = \left(\sum_{i=1}^{\text{rk}(A)} \sigma_i(A)^p\right)^{1/p}$$
A few matrix norms

Let $A$ be a matrix. The singular values of $A$ are $\sigma_i(A) = \sqrt{\lambda_i(AA^T)}$.

Define

$$\|A\|_p = \ell_p(\sigma) = \left( \sum_{i=1}^{\text{rk}(A)} \sigma_i(A)^p \right)^{1/p}$$

- Trace norm: $\|A\|_1$
A few matrix norms

Let $A$ be a matrix. The singular values of $A$ are $\sigma_i(A) = \sqrt{\lambda_i(\mathbf{AA}^T)}$.

Define

$$ \|A\|_p = \ell_p(\sigma) = \left( \sum_{i=1}^{\text{rk}(A)} \sigma_i(A)^p \right)^{1/p} $$

- **Trace norm:** $\|A\|_1$
- **Spectral norm:** $\|A\|_{\infty}$
A few matrix norms

Let $A$ be a matrix. The singular values of $A$ are $\sigma_i(A) = \sqrt{\lambda_i(AA^T)}$.

Define

$$\|A\|_p = \ell_p(\sigma) = \left( \sum_{i=1}^{\text{rk}(A)} \sigma_i(A)^p \right)^{1/p}$$

- Trace norm: $\|A\|_1$
- Spectral norm: $\|A\|_\infty$
- Frobenius norm $\|A\|_2$
A few matrix norms

Let $A$ be a matrix. The singular values of $A$ are $\sigma_i(A) = \sqrt{\lambda_i(AA^T)}$. Define

$$\|A\|_p = \ell_p(\sigma) = \left( \sum_{i=1}^{\text{rk}(A)} \sigma_i(A)^p \right)^{1/p}$$

- Trace norm: $\|A\|_1$
- Spectral norm: $\|A\|_\infty$
- Frobenius norm $\|A\|_2 = \sqrt{\text{Tr}(AA^T)}$
A few matrix norms

Let $A$ be a matrix. The singular values of $A$ are $\sigma_i(A) = \sqrt{\lambda_i(AA^T)}$.

Define

$$\|A\|_p = \ell_p(\sigma) = \left( \sum_{i=1}^{\text{rk}(A)} \sigma_i^p(A) \right)^{1/p}$$

- Trace norm: $\|A\|_1$
- Spectral norm: $\|A\|_\infty$
- Frobenius norm $\|A\|_2 = \sqrt{\text{Tr}(AA^T)} = \sqrt{\sum_{i,j} |A_{ij}|^2}$
Example: trace norm

As $\ell_1$ and $\ell_\infty$ are dual, so too are trace norm and spectral norm:

$$\| A \|_1 = \max_B \frac{|\langle A, B \rangle|}{\| B \|_\infty}$$
Example: trace norm

As $\ell_1$ and $\ell_\infty$ are dual, so too are trace norm and spectral norm:

$$\|A\|_1 = \max_B \frac{|\langle A, B \rangle|}{\|B\|_\infty}$$

- Thus to show that the trace norm of $A$ is large, it suffices to find $B$ with non-negligible inner product with $A$ and small spectral norm.
Example: trace norm

As $\ell_1$ and $\ell_\infty$ are dual, so too are trace norm and spectral norm:

$$\|A\|_1 = \max_B \frac{|\langle A, B \rangle|}{\|B\|_\infty}$$

- Thus to show that the trace norm of $A$ is large, it suffices to find $B$ with non-negligible inner product with $A$ and small spectral norm.
- We will refer to $B$ as a *witness*. 
Application to communication complexity

- For a function $f : X \times Y \rightarrow \{-1,+1\}$ we define the communication matrix $A_f[x,y] = f(x,y)$.

- For deterministic communication complexity, one of the best lower bounds available is log rank:
  $$D(f) \geq \log \text{rk}(A_f)$$

- The famous log rank conjecture states this lower bound is polynomially tight.
Application to communication complexity

• As the rank of a matrix is equal to the number of non-zero singular values, we have

\[ \| A \|_1 = \sum_{i=1}^{\text{rk}(A)} \sigma_i(A) \]
Application to communication complexity

As the rank of a matrix is equal to the number of non-zero singular values, we have

\[
\|A\|_1 = \sum_{i=1}^{\text{rk}(A)} \sigma_i(A) \leq \sqrt{\text{rk}(A)} \left( \sum_i \sigma_i^2(A) \right)^{1/2}
\]
Application to communication complexity

• As the rank of a matrix is equal to the number of non-zero singular values, we have

\[ \|A\|_1 = \sum_{i=1}^{\text{rk}(A)} \sigma_i(A) \leq \sqrt{\text{rk}(A)} \left( \sum_i \sigma_i^2(A) \right)^{1/2} = \sqrt{\text{rk}(A)} \|A\|_2 \]
Application to communication complexity

- As the rank of a matrix is equal to the number of non-zero singular values, we have

\[ \|A\|_1 = \sum_{i=1}^{\text{rk}(A)} \sigma_i(A) \leq \sqrt{\text{rk}(A)} \left( \sum_i \sigma_i^2(A) \right)^{1/2} = \sqrt{\text{rk}(A)} \|A\|_2 \]

- For a $M$-by-$N$ sign matrix $\|A\|_2 = \sqrt{MN}$ so we have

\[ 2^{D(f)} \geq \text{rk}(A_f) \geq \frac{(\|A_f\|_1)^2}{MN} \]

Call this the “trace norm method.”
Trace norm method (example)

• Let $H_N$ be a $N$-by-$N$ Hadamard matrix (entries from $\{-1, +1\}$).

• Then $\|H_N\|_1 = N^{3/2}$.

• Trace norm method gives bound on rank of $N^3/N^2 = N$
Trace norm method (drawback)

- As a complexity measure, the trace norm method suffers one drawback—it is not monotone.
  \[
  \begin{pmatrix}
  H_N & 1_N \\
  1_N & 1_N
  \end{pmatrix}
  \]

- Trace norm at most \( N^{3/2} + 3N \)
Trace norm method (drawback)

• As a complexity measure, the trace norm method suffers one drawback—it is not monotone.

\[
\begin{pmatrix}
H_N & 1_N \\
1_N & 1_N
\end{pmatrix}
\]

• Trace norm at most \( N^{3/2} + 3N \)

• Trace norm method gives

\[
\frac{(N^{3/2} + 3N)^2}{4N^2}
\]

worse bound on whole than on \( H_N \) submatrix!
Trace norm method (a fix)

• We can fix this by considering

$$\max_{u,v:} \| A \circ uv^T \|_1$$

$$\|u\|_2 = \|v\|_2 = 1$$
Trace norm method (a fix)

• We can fix this by considering

$$\max_{u,v: \|u\|_2 = \|v\|_2 = 1} \| A \circ uv^T \|_1$$

• As $\text{rk}(A \circ uv^T) \leq \text{rk}(A)$ we still have

$$\text{rk}(A) \geq \left( \frac{\| A \circ uv^T \|_1}{\| A \circ uv^T \|_2} \right)^2$$
The $\gamma_2$ norm

- This bound simplifies nicely for a sign matrix $A$

$$\text{rk}(A) \geq \max_{u,v: \|u\|_2=\|v\|_2=1} \left( \frac{\|A \odot uv^T\|_1}{\|A \odot uv^T\|_2} \right)^2 \frac{\|u\|_2 = \|v\|_2 = 1}{\|A \odot uv^T\|_2}$$
The $\gamma_2$ norm

- This bound simplifies nicely for a sign matrix $A$

$$\text{rk}(A) \geq \max_{u,v: \|u\|_2 = \|v\|_2 = 1} \left( \frac{\|A \circ uv^T\|_1}{\|A \circ uv^T\|_2} \right)^2 = \max_{u,v: \|u\|_2 = \|v\|_2 = 1} \|A \circ uv^T\|_1$$

- We have arrived at the $\gamma_2$ norm introduced to communication complexity by [LMSS07, LS07]

$$\gamma_2(A) = \max_{u,v: \|u\|_2 = \|v\|_2 = 1} \|A \circ uv^T\|_1$$
\( \gamma_2 \) norm: Surprising usefulness

- In matrix analysis known as “Schur/Hadamard product operator/trace norm,”
\( \gamma_2 \) norm: Surprising usefulness

- In matrix analysis known as “Schur/Hadamard product operator/trace norm,”

- Schur (1911) showed that \( \gamma_2(A) = \max_i A_{ii} \) if \( A \) psd.
\( \gamma_2 \) norm: Surprising usefulness

- In matrix analysis known as “Schur/Hadamard product operator/trace norm,”

- Schur (1911) showed that \( \gamma_2(A) = \max_i A_{ii} \) if \( A \) psd.

- The dual norm \( \gamma_2^*(A) = \max_B \langle A, B \rangle / \gamma_2(B) \) turns up in semidefinite programming relaxation of MAX-CUT of Goemans and Williamson, and quantum value of two-player XOR games.
\( \gamma_2 \) norm: Surprising usefulness

- In matrix analysis known as “Schur/Hadamard product operator/trace norm,”

- Schur (1911) showed that \( \gamma_2(A) = \max_i A_{ii} \) if \( A \) psd.

- The dual norm \( \gamma_2^*(A) = \max_B \langle A, B \rangle / \gamma_2(B) \) turns up in semidefinite programming relaxation of MAX-CUT of Goemans and Williamson, and quantum value of two-player XOR games

- \( \text{disc}_P(A) = \Theta(\gamma_2^*(A \circ P)) \) [Linial Shraibman 08]
Randomized and quantum communication complexity

• So far it is not clear how much we have gained. Many techniques available to bound matrix rank.

• But for randomized and quantum communication complexity the relevant measure is no longer rank, but *approximation rank* [BW01]. For a sign matrix $A$:

$$ \text{rk}_\alpha(A) = \min_B \{ \text{rk}(B) : 1 \leq A[x, y] \cdot B[x, y] \leq \alpha \} $$

• Limiting case is sign rank: $\text{rk}_\infty(A) = \min \{ \text{rk}(B) : 1 \leq A[x, y] \circ B[x, y] \}$. 

• NP-hard? Can be difficult even for basic matrices.
Approximation norms

• We have seen how trace norm and $\gamma_2$ lower bound rank.

• In a similar fashion to approximation rank, we can define approximation norms. For an arbitrary norm $||| \cdot |||$ let

$$|||A|||^\alpha = \min_{B} \{|||B||| : 1 \leq A[x, y] \cdot B[x, y] \leq \alpha\}$$

• Note that an approximation norm is not itself necessarily a norm.

• However, we can still use duality to obtain a max expression

$$|||A|||^\alpha = \max_{B} \frac{(1 + \alpha)\langle A, B \rangle + (1 - \alpha)\ell_1(B)}{2|||B|||^*}$$
From our discussion, for a $M$-by-$N$ sign matrix $A$

$$\text{rk}_\alpha(A) \geq \frac{\gamma_2^\alpha(A)^2}{\alpha^2} \geq \frac{(\|A\|_1^\alpha)^2}{\alpha^2 MN}$$
Approximate $\gamma_2$

- From our discussion, for a \( M \)-by-\( N \) sign matrix \( A \)

\[
\text{rk}_\alpha(A) \geq \frac{\gamma_2^\alpha(A)^2}{\alpha^2} \geq \frac{\|A\|_1^\alpha}{\alpha^2 MN}
\]

- We show that for any sign matrix \( A \) and constant \( \alpha > 1 \)

\[
\text{rk}_\alpha(A) = O \left( \gamma_2^\alpha(A)^2 \log(MN) \right)^3
\]
Remarks

- When $\alpha = 1$ theorem does not hold. For equality function (sign matrix) $\text{rk}(2I_N - 1_N) \geq N - 1$, but

$$\gamma_2(2I_N - 1_N) \leq 2\gamma_2(I_N) + \gamma_2(1_N) = 3,$$

by Schur’s theorem.

- Equality example also shows that the $\log N$ factor is necessary, as approximation rank of identity matrix is $\Omega(\log N)$ [Alon 08].
Advantages of $\gamma_2^\alpha$

- $\gamma_2^\alpha$ can be formulated as a max expression

$$\gamma_2^\alpha(A) = \max_B \frac{(1 + \alpha)\langle A, B \rangle + (1 - \alpha)\ell_1(B)}{2\gamma_2^*(B)}$$

- $\gamma_2^\alpha$ is polynomial time computable by semidefinite programming

- $\gamma_2^\alpha$ is also known to lower bound quantum communication with shared entanglement, which was not known for approximation rank.
Proof sketch

- A dual formulation of trace norm $\|A\|_1 = \min_{X,Y: \quad X^T Y = A} \|X\|_2 \|Y\|_2$
Proof sketch

• A dual formulation of trace norm \( \| A \|_1 = \min_{X,Y: \ X^TY=A} \| X \|_2 \| Y \|_2 \)

• Similarly, \( \gamma_2 \) has the min formulation

\[
\gamma_2(A) = \min_{X,Y: \ X^TY=A} c(X)c(Y)
\]

where \( c(X) \) is the maximum \( \ell_2 \) norm of a column of \( X \).
Proof sketch

• A dual formulation of trace norm $\|A\|_1 = \min_{X,Y: X^TY = A} \|X\|_2 \|Y\|_2$

• Similarly, $\gamma_2$ has the min formulation

$$\gamma_2(A) = \min_{X,Y: X^TY = A} c(X)c(Y)$$

where $c(X)$ is the maximum $\ell_2$ norm of a column of $X$.

• Rank can also be phrased as optimizing over factorizations: the minimum $K$ such that $A = X^TY$ where $X, Y$ are $K$-by-$N$ matrices.
First step: dimension reduction

- Look at $X^T Y = A'$ factorization realizing $\gamma_2^{1+\epsilon}(A)$. Say $X, Y$ are $K$-by-$N$. 
First step: dimension reduction

- Look at $X^T Y = A'$ factorization realizing $\gamma_2^{1+\epsilon}(A)$. Say $X, Y$ are $K$-by-$N$.

- Know that the columns of $X, Y$ have squared $\ell_2$ norm at most $\gamma_2(A')$, but $X, Y$ might still have many rows...
First step: dimension reduction

- Look at $X^T Y = A'$ factorization realizing $\gamma_2^{1+\epsilon}(A)$. Say $X, Y$ are $K$-by-$N$.

- Know that the columns of $X, Y$ have squared $\ell_2$ norm at most $\gamma_2(A')$, but $X, Y$ might still have many rows...

- Consider $RX$ and $RY$ where $R$ is random matrix of size $K'$-by-$K$ for $K' = O(\gamma_2^{1+\epsilon}(A)^2 \log N)$. By Johnson-Lindenstrauss lemma whp all the inner products $(RX)_i^T (RY)_j \approx X_i^T Y_j$ will be approximately preserved.
First step: dimension reduction

• Look at $X^TY = A'$ factorization realizing $\gamma_2^{1+\epsilon}(A)$. Say $X, Y$ are $K$-by-$N$.

• Know that the columns of $X, Y$ have squared $\ell_2$ norm at most $\gamma_2(A')$, but $X, Y$ might still have many rows...

• Consider $RX$ and $RY$ where $R$ is random matrix of size $K'$-by-$K$ for $K' = O(\gamma_2^{1+\epsilon}(A)^2 \log N)$. By Johnson-Lindenstrauss lemma whp all the inner products $(RX)_i^T(RY)_j \approx X_i^TY_j$ will be approximately preserved.

• This shows there is a matrix $A'' = (RX)^T(RY)$ which is a $1 + 2\epsilon$ approximation to $A$ and has rank $O(\gamma_2^{1+\epsilon}(A)^2 \log N)$. 
Second step: Error reduction

• Now we have a matrix $A'' = (RX)^T(RY)$ which is of the desired rank, but is only a $1 + 2\epsilon$ approximation to $A$, whereas we wanted an $1 + \epsilon$ approximation of $A$.

• Idea [Alon 08, Klivans Sherstov 07]: apply a polynomial to the entries of the matrix. Can show $\text{rk}(p(A)) \leq (d+1)\text{rk}(A)^d$ for degree $d$ polynomial.

• Taking $p$ to be low degree approximation of sign function makes $p(A'')$ better approximation of $A$. For our purposes, can get by with degree 3 polynomial.

• Completes the proof $\text{rk}_\alpha(A) = O\left(\gamma_2^\alpha(A)^2 \log(N)\right)^3$
Polynomial for Error Reduction

\[ \frac{7x - x^3}{6} \]
Multiparty complexity: Number on the forehead model

- Now we have $k$-players and a function $f : X_1 \times \ldots \times X_k \rightarrow \{-1, +1\}$. Player $i$ knows the entire input except $x_i$.

- This model is the “frontier” of communication complexity. Lower bounds have nice applications to circuit and proof complexity.

- Instead of communication matrix, have communication tensor $A_f[x_1, \ldots, x_k] = f(x_1, \ldots, x_k)$. This makes extension of linear algebraic techniques from the two-party case difficult.

- Only method known for general model of number-on-the-forehead is discrepancy method.
Discrepancy method

• Two-party case
  \[ \text{disc}_P(A) = \max_C \langle A \circ P, C \rangle \]
  where \( C \) is a combinatorial rectangle.

• For NOF model, analog of combinatorial rectangle is cylinder intersection. For a tensor \( A \),
  \[ \text{disc}_P(A) = \max_C \langle A \circ P, C \rangle \]
  where \( C \) is a cylinder intersection.

• In both cases,
  \[ R_\epsilon(A) \geq \max_P \frac{1 - 2\epsilon}{\text{disc}_P(A)} \]
Discrepancy method

- For some functions like generalized inner product, discrepancy can show nearly optimal bounds $\Omega(n/2^{2k})$ [BNS89]

- But for other functions, like disjointness, discrepancy can only show lower bounds $O(\log n)$. Follows as discrepancy actually lower bounds non-deterministic complexity.

- The best lower bounds on disjointness were $\Omega(\frac{\log n}{k})$ [T02, BPSW06].
Norms for multiparty complexity

• Basic fact: A successful $c$-bit NOF protocol partitions the communication tensor into at most $2^c$ many monochromatic cylinder intersections.

• This allows us to define a norm

$$\mu(A) = \min \{ \sum |\gamma_i| : A = \sum \gamma_i C_i \}$$

$C_i$ is a cylinder intersection.

• We have $D(A) \geq \log \mu(A)$. For matrices $\mu(A) = \Theta(\gamma_2(A))$
Norms for multiparty complexity

• Basic fact: A successful $c$-bit NOF protocol partitions the communication tensor into at most $2^c$ many monochromatic cylinder intersections.

• This allows us to define a norm

$$\mu(A) = \min\left\{ \sum |\gamma_i| : A = \sum \gamma_i C_i \right\}$$

$C_i$ is a cylinder intersection.

• We have $D(A) \geq \log \mu(A)$. For matrices $\mu(A) = \Theta(\gamma_2(A))$

• Also, by usual arguments get $R_{\epsilon}(A) \geq \mu^{\alpha}(A)$ for $\alpha = 1/(1 - 2\epsilon)$. 
The dual norm

The dual norm is

$$\mu^*(A) = \max_{B: \mu(B) \leq 1} |\langle A, B \rangle|$$
The dual norm

- The dual norm is
  \[ \mu^*(A) = \max_{B: \mu(B) \leq 1} |\langle A, B \rangle| \]

- So we see \( \mu^*(A) = \max_C |\langle A, C \rangle| \) where \( C \) is a cylinder intersection.
The dual norm

• The dual norm is

$$\mu^*(A) = \max_{B : \mu(B) \leq 1} |\langle A, B \rangle|$$

• So we see $\mu^*(A) = \max_C |\langle A, C \rangle|$ where $C$ is a cylinder intersection.

• $\text{disc}_P(A) = \mu^*(A \circ P)$
The dual norm

• The dual norm is

\[ \mu^*(A) = \max_{B: \mu(B) \leq 1} |\langle A, B \rangle| \]

• So we see \( \mu^*(A) = \max_{C} |\langle A, C \rangle| \) where \( C \) is a cylinder intersection.

• \( \text{disc}_P(A) = \mu^*(A \circ P) \)

• Bound \( \mu^\alpha(A) \) in the following form. Standard discrepancy is exactly \( \mu^\infty(A) \).

\[ \mu^\alpha(A) = \max_B \frac{(1 + \alpha)\langle A, B \rangle + (1 - \alpha)\ell_1(B)}{2\mu^*(B)} \]
A limiting case

• Recall the bound

$$\mu^\alpha(A) = \max_B \frac{(1 + \alpha) \langle A, B \rangle + (1 - \alpha) \ell_1(B)}{2\mu^*(B)}$$

• As $\alpha \to \infty$, larger penalties for entries where $B[x,y]$ differs in sign from $A[x,y]$

$$\mu^\infty(A) = \max_{B : A \odot B \geq 0} \frac{\langle A, B \rangle}{\mu^*(B)}$$

• This is just the standard discrepancy method.
Choosing a witness

- As $\mu^\alpha$ is decreasing function of $\alpha$, we have a technique stronger than the discrepancy method.
Choosing a witness

- As \( \mu^\alpha \) is decreasing function of \( \alpha \), we have a technique stronger than the discrepancy method. Yet choosing a witness is still a difficult problem.
Choosing a witness

• As $\mu^{\alpha}$ is decreasing function of $\alpha$, we have a technique stronger than the discrepancy method. Yet choosing a witness is still a difficult problem.

• Use framework of pattern matrices [Sherstov 07, 08] and generalization to pattern tensors in multiparty case [Chattopadhyay 07]: For functions of the form $f(x_1 \land \ldots \land x_k)$, can choose witness derived from dual polynomial witnessing that $f$ has high approximate degree.

• Degree/Discrepancy Theorem [Sherstov 07,08 Chattopadhyay 08]: Pattern tensor derived from function with pure high degree will have small discrepancy. In multiparty case, this uses [BNS 89] technique of bounding discrepancy.
Final result

- Final result: Randomized $k$-party complexity of disjointness

$$\Omega \left( \frac{n^{1/(k+1)}}{2^{2^k}} \right)$$

- Independently shown by Chattopadhyay and Ada

- Beame and Huynh-Ngoc have recently shown non-trivial lower bounds on disjointness for up to $\log^{1/3} n$ players (though not as strong as ours for small $k$).
An open question

• We have shown a polynomial time algorithm to approximate $\text{rk}_\alpha(A)$, but ratio deteriorates as $\alpha \to \infty$.

$$\frac{\gamma_2^\alpha(A)^2}{\alpha^2} \leq \text{rk}_\alpha(A) \leq O \left( \gamma_2^\alpha(A)^2 \log(N) \right)^3$$

• For the case of sign rank, lower bound fails! In fact, exponential gaps are known [BVW07, Sherstov07]

• Polynomial time algorithm to approximate sign rank?